Machine learning and deep learning approaches for multivariate time series prediction and anomaly detection

Thill, M.

Machine Learning and Deep Learning Approaches for Multivariate Time Series Prediction and Anomaly Detection

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 17 maart 2022
klokke 16.15 uur

door

MARKUS THILL

geboren te Patan, Nepal
in 1987

Promotiecommissie

Promotores:	Prof.dr. T.H.W. Bäck	
	Prof.dr. W. Konen	(Cologne University of Applied Sciences)
Co-promotor:	Dr. H. Wang	
Promotiecommissie:	Prof.dr. A. Plaat	
	Prof.dr. K.J. Batenburg	
	Dr. M. Baratchi	
	Prof.dr. H. Trautmann	(University of Münster)
	Prof.dr. E. Alba	(University of Malaga)

Abstract

In many real-world applications today, it is critical to continuously record and monitor certain machine or system health indicators in order to discover malfunctions or other abnormal behavior at an early stage and prevent potential harm. The demand for such reliable monitoring systems is expected to increase in the coming years. Particularly in the industrial context, in the course of ongoing digitization, it is becoming increasingly important to analyze growing volumes of data in an automated manner using state-of-the-art algorithms. In many practical applications, one has to deal with temporal data in the form of data streams or time series. The problem of detecting unusual (or anomalous) behavior in time series is commonly referred to as time series anomaly detection. Anomalies are events observed in the data that do not conform to the normal or expected behavior when viewed in their temporal context.

In the era of 'industry 4.0', 'cyber-physical systems' and 'big data', data-driven AI (artificial intelligence) approaches (designed to learn solely through data) from the fields of machine learning (ML), or deep learning (DL) – a sub-field of ML – have gained tremendous popularity. The majority of ML models are trained in a supervised fashion and require a labeled training data set. However, a particular problem in (time series) anomaly detection is that labeled data are usually relatively sparse and, as a consequence, supervised learning methods are mostly not feasible.

This thesis focuses on unsupervised machine learning algorithms for anomaly detection in time series. In an unsupervised learning setup, a model attempts to learn the normal behavior in a time series – which might already be contaminated with anomalies – without any external assistance. The model can then use its learned notion of normality to detect anomalous events. This work presents four unsupervised anomaly detection algorithms for multivariate time series, which can be used in different contexts: 1. SORAD learns an autoregressive prediction model and maintains a distribution of the prediction errors to detect anomalies. It can operate fully online and is up-and-running after a very short transient phase. Due to its online adaptability, it is advantageous for non-stationary time series. 2. DWT-MLEAD uses discrete wavelet transforms (DWT) to analyze a time series signal at different frequency scales to detect short-range and longer-range anomalies. It is also online-adaptable and computationally efficient. 3. LSTM-AD is a DL approach designed for longer time series, which uses a stack of recurrent long short-term memory (LSTM) neural networks to predict normal time-series behavior. Due to several extensions, it is particularly well-suited for quasi-periodic time series. 4. TCN-AE is a deep temporal autoencoder architecture, which – similarly to DWT-MLEAD – analyzes time series at different frequency scales and can learn short- and long-term relationships. It uses so-called dilated convolutional layers with learnable filters. TCN-AE is usually more compact in terms of model size than other DL models and can be trained faster due to its convolutional architecture.

We evaluate the algorithms presented in this work on challenging synthetic and real-world time series anomaly detection benchmarks, such as Mackey-Glass time series or electrocardiogram recordings, and compare them to other state-of-the-art algorithms.

Contents

List of Tables

List of Figures